

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2001, 18:42:19 ; Search time 2080.33 Seconds
(without alignments)
9549.392 Million cell updates/sec

Title: US-09-494-297-1
Perfect score: 2274
Sequence: 1 atgaaaaaaaaggttcc.....ggataagaacaatgactag 2274

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthun10:*
45: em_esthun11:*
46: em_esthun12:*
47: em_esthun13:*
48: em_esthun14:*
49: em_esthun15:*
50: em_esthun16:*
51: em_esthun17:*
52: em_esthun18:*
53: em_esthun19:*
54: em_esthun20:*
55: em_esthun21:*
56: em_esthun22:*
57: em_esthun23:*
58: em_esthun24:*
59: em_esthun25:*
60: em_esthun26:*
61: em_esthun27:*
62: em_esthun28:*
63: em_esthun29:*
64: em_esthun30:*
65: em_esthun31:*
66: em_esthun32:*
67: em_esthun33:*
68: em_esthun34:*
69: em_esthun35:*
70: em_esthun36:*
71: em_esthun37:*
72: em_esthun38:*
73: em_esthun39:*
74: em_esthun40:*
75: em_esthun41:*
76: em_esthun42:*
77: em_esthun43:*
78: em_esthun44:*
79: em_esthun45:*
80: em_esthun46:*
81: em_esthun47:*
82: em_esthun48:*
83: em_esthun49:*
84: em_esthun50:*
85: em_esthun51:*
86: em_esthun52:*
87: em_esthun53:*
88: em_esthun54:*
89: em_esthun55:*
90: em_esthun56:*
91: em_esthun57:*
92: em_esthun58:*
93: em_esthun59:*
94: em_esthun60:*
95: em_esthun61:*
96: em_esthun62:*
97: em_esthun63:*
98: em_esthun64:*
99: em_esthun65:*
100: em_esthun66:*
101: em_esthun67:*
102: em_esthun68:*
103: em_esthun69:*
104: em_esthun70:*
105: em_esthun71:*
106: em_esthun72:*
107: em_esthun73:*
108: em_esthun74:*
109: em_esthun75:*
110: em_esthun76:*
111: em_esthun77:*
112: em_esthun78:*
113: em_esthun79:*
114: em_esthun80:*
115: em_esthun81:*
116: em_esthun82:*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

190: em_gss_pln1:*
 191: em_gss_pln2:*
 192: em_gss_pro:*
 193: em_gss_rod1:*
 194: em_gss_rod2:*
 195: em_gss_rod3:*
 196: em_gss_rod4:*
 197: em_gss_rod5:*
 198: em_gss_vrt1:*
 199: em_gss_vrt2:*
 200: em_gss_vrt3:*
 201: gb_gss1:*
 202: gb_gss2:*
 203: gb_gss3:*
 204: gb_gss4:*
 205: gb_gss5:*
 206: gb_gss6:*
 207: gb_gss7:*
 208: gb_gss8:*
 209: gb_gss9:*
 210: gb_gss10:*
 211: gb_gss11:*
 212: gb_gss12:*
 213: gb_gss13:*
 214: gb_gss14:*
 215: gb_gss15:*
 216: gb_gss16:*
 217: gb_gss17:*
 218: gb_gss18:*
 219: gb_gss19:*
 220: gb_gss20:*
 221: gb_gss21:*
 222: gb_gss22:*
 223: gb_gss23:*
 224: gb_gss24:*
 225: gb_gss25:*
 226: gb_gss26:*
 227: gb_gss27:*
 228: gb_gss28:*
 229: gb_gss29:*
 230: gb_gss30:*
 231: gb_gss31:*
 232: gb_gss32:*
 233: gb_gss33:*
 234: gb_gss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	70.2	3.1	1101	229	CNS0039G	AL063921 Drosophila
2	61.4	2.7	1101	229	CNS00EVL	AL069706 Drosophila
3	61	2.7	1201	229	CNS0167M	AL106396 Drosophila
4	60.2	2.6	823	225	AZ687380	AZ687380 EMTL056TF
5	60.2	2.6	1092	230	CNS020K7	AL175696 Tetradodon
6	60.2	2.6	1101	229	CNS01219	AL101595 Drosophila
7	60	2.6	1101	229	CNS00EPO	AL064493 Drosophila
8	59.2	2.6	1032	230	CNS020IP	AL206746 Tetradodon
9	58.6	2.6	1200	229	CNS016CO	AL106578 Drosophila
10	58	2.6	668	105	AL514901	AL514901 AL514901
11	57.8	2.5	891	225	AZ683582	AZ683582 EMTK47TR
12	57.4	2.5	652	231	CNS03HOU	AL244551 Tetradodon
13	57.4	2.5	1101	229	CNS000B8	AL063632 Drosophila
14	56.8	2.5	1101	229	CNS00KAE	AL077628 Drosophila
15	56.4	2.5	928	146	BF274633	BF274633 GA_EB002
16	56.2	2.5	881	223	AZ547573	AZ547573 EMTG88TR
17	56	2.5	1101	229	CNS00EO7	AL069440 Drosophila
18	55.8	2.5	611	201	AA549883	AA549883 0956m3 gm

19	55.8	2.5	877	105	AL514453	AL514453	AL514453
20	55.8	2.5	884	225	A2673406	A2673406	ENTMR22TFF
C 21	55.8	2.5	1101	229	CNSO039L	AL063926	Drosophil
C 22	55.6	2.4	900	229	CNSO1534	AL105514	Drosophil
C 23	55.4	2.4	636	202	AO159098	AO159098	nbxb0012P
C 24	55.2	2.4	828	229	CNSO11TX	AL100719	Drosophil
C 25	55.2	2.4	1025	229	CNSO14J2	AL104216	Drosophil
26	54.8	2.4	908	223	A2548467	A2548467	ENTER30TFF
27	54.8	2.4	1001	229	CNSO155H	AL105023	Drosophil
28	54.4	2.4	750	229	CNSO11ID	AL100303	Drosophil
29	54.2	2.4	734	229	CNSO10MP	AL099163	Drosophil
30	54	2.4	767	229	CNSO08OX	AL055924	Drosophil
31	54	2.4	928	229	CNSO0DKY	AL071865	Drosophil
C 32	54	2.4	982	205	AO325799	AQ325799	nbxb0021B
C 33	53.8	2.4	843	223	A2551618	A2551618	ENTDV54TFR
C 34	53.6	2.4	640	18	A1258374	A1258374	LP01554.5
C 35	53.6	2.4	877	223	A2545974	A2545974	ENTER32TFF
C 36	53.4	2.3	669	230	CNSO3C2I	AL227267	Tetradodon
C 37	53.4	2.3	757	105	AL514421	AL514421	AL514421
C 38	53.4	2.3	928	229	CNSO0DKY	AL071865	Drosophil
C 39	53.4	2.3	1101	229	CNSO0DTW	AL069847	Drosophil
C 40	53.2	2.3	994	231	CNSO4NOJ	AL286972	Tetradodon
C 41	53	2.3	834	227	B12387	B12387	F21E20-Sp6
C 42	53	2.3	865	225	A2674459	A2674459	ENTIRP90TFF
C 43	53	2.3	950	225	A2672116	A2672116	ENTFEM06TFF
C 44	52.8	2.3	761	230	CNSO2MEY	AL204019	Tetradodon
C 45	52.8	2.3	905	223	A2550256	A2550256	ENTEW58TFR

ALIGNMENTS

	BASE COUNT	201 a	/note="end : TET3" 64 C	131 g	202 t	503 others
OY	Query Match Best Local Similarity Matches 121; Conservative	3.1%; 17.5% 286;	Score 70.2; Pred. No. 1,8e-05; Mismatches 283;	DB 229; Length 1101; Indels 1;	Gaps 1	
OY	107 tgatcttcgcttgtaacttccatggtttggcgtcaagcctlttttggtttaagaagt 166	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Dd	1091 WBRDRKDDMDWTKKWTTWKWKRRADRRAGDADRANWDGAGTWTATWWMMWMTATMD 1032	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
OY	167 cctcgacgcgaacaataatccagattcaagttcggaaatcacagaytgtatbtgatcy 226	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Dd	1031 TWMDKMMWMTAAKTIDTAWTMTWTRAFRADWAGRGCRDRRDAATDADGAGRDRGRKR 972	:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
OY	227 aatcttatgtaaaggagcatccatatataaacagttttgagttagcacacgatttaaagg 286	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Dd	971 KDKRKRDDEDDDKKGSKKKAARKAKMWATKMWDMWMDKDWMWGAKDKRRADDDEGADGK 912	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
OY	287 ttactatgaagaagtagaagttatccaagttatcgtttaattaaagaacattcc 346	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
Dd	911 DDDGGKDDDDDTGTGTDKDDDDKDKWMDMKAGCTWGDATWANAATIDMMWGNADADPMWTW 852	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				
OY	347 ctctcggatcacagatagttgtlaaaaagtygtlataaaaaacatgacatggaatctctaca 406	::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::				

RESULT	1
LOCUS	CNS00396/c
DEFINITION	CNS00396 1101 bp DNA GSS 03-JUN-1999
ACCESSION	Drosophila melanogaster genome survey sequence TERT3 end of BAC #
VERSION	BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
SOURCE	fly), genomic survey sequence.
KEYWORDS	AT063921
ORGANISM	AT063921.1 GI:4941778
REFERENCE	GSS.
AUTHORS	fruit fly.
TITLE	Drosophila melanogaster
JOURNAL	Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Prerygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
COMMENT	1 (bases 1 to 1101) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 101 91006 Evry cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aaron Mammaster in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm . location/Qualifiers 1. 1101 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_1lb="RPCI-98" /clone="BACR08K10"

[illegible]

Db 918 AAAAAAAAAAATTTTTTTTAAAAAAATTAATAAATTATTAATAAATAAATAA 859

Oy 1736 gaactcagtcgaccccaagaattagtcatatctcgtaatgaagataaaaaaag 1795
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 858 AAAAAAAAAAATTTTTTTTAAAAAAATTAATAAATTATTAATAAATAAATAA 799

Oy 1796 ttataccgttaaccabatttaacatctgagaaaacggtagctggtttagctgtgca 1855
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 798 TTTTNTNAAAATTAATAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 739

Oy 1856 gaactaaagattccatttgaaatgaaatgaataaaaaataacaagaattgcttc 1915
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 738 ACAAAAAAATTTTAAATTAATAAATTAATAAATTAATAAATTAATAAATTTT 679

Oy 1916 aaactgttaaacaagataaacacaacctgcgaattaaagatgtaagcaacatt 1975
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 678 MAMMAMMMNNNNCCAAAAAAMMAAVVAAAACMAMMTTAAAGCTTAAAAAACMAA 619

Oy 1976 taaacatggggaaagtttaacactcaagggttacccaagaagttatctta 2035
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 618 MAMMAMMTTAAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 559

Oy 2036 aagaaa 2041

Db 558 CABANA 553

RESULT 4

AZ687380 823 bp DNA GSS 14-DEC-2000

LOCUS ENTLOS6TF Entamoeba histolytica Sheared DNA Entamoeba histolytica

DEFINITION genomic, DNA sequence.

VERSION A2687380 GI:11824526

KEYWORDS GSS.

SOURCE Entamoeba histolytica.

ORGANISM Entamoeba histolytica.
Eukaryota; Entamoebidae; Entamoeba.
1 (bases 1 to 823)
Loftus,B., Van Aken,S. and Fraser,C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)

JOURNAL Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: b.loftus@iglr.org

COMMENT Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: ML3-Forward
Class: Shotgun
High quality sequence start: 22
High quality sequence stop: 820.
Location/Qualifiers
1..823
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_id="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOSt1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, U.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome

BASE COUNT	400 a	81 c	163 g	179 t	
ORIGIN	<p>Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).</p>				
Query Match	2.6%; Score 60.2; DB 225; Length 823;				
Best Local Similarity	45.9%; Pred. No. 0.003; Indels 9; Gaps 2				
Matches 286; Conservative	0; Mismatches 328; Indels 9; Gaps 2				
Qy 1583	aaccaaaagactatcatggttttgagacatbaatgatgatcttaagcaatgtctctaa 1642				
Db 67	AAACACATTTGAAACCTTGCGTGAACGAGAAAGAGCTTGCTCTCAATGCTCTACTA 126				
Qy 1643	tcctttagaatacgcctcaagatagtaatcctccacagctaacctgacttattcttta 1702				
Db 127	AGAAAGCATCTGATGAAAGAGATCTTTAACTCAAAATCTTGCAANTGAGAAATTACTA 186				
Qy 1703	ttccgaataacaataataatcatcctctattggaaacctcagtgacatccagaatttag 1762				
Db 187	CTAAGAAATCTTACTAAAAACAAAGCAGATTTGGAAAAAGATTAAGTGAATTAACAAAG 246				
Qy 1763	ttgatattatctgfatgagaataaaaaagattatcaccttcaactatataatttaacat 1822				
Db 247	ATTATGACAGTTTATGACAGATGATTAAGA-----TAAGTTTAAGCTGATTTTGACAAATG 300				
Qy 1823	tgagaaaaacggtgactggtttagctgtgtgacagaactaaagatttccatttgaattg 1882				
Db 301	CTCAAGAAAGATTAAGAAAGCTTGATGATGAATTAATCTTAAGACAGACAGATGTTTCACAA 360				
Qy 1883	aatataaaaaataatagcagaagatttgccttccaaactggttaaacagatataaacaaac 1942				
Db 361	ATCTTCAAAAACAAAAAGAGAAATATGAATCAAAATTCCTAAATGCAAGAGAAAAAGG 420				
Qy 1943	tcgaattt---aaagatgfaaagcaacatcaatttaaacatgvggaaagttaacac 1999				
Db 421	AAGCAATTTGCAATGATGATTTAAAAATTAAGAAAAAGCAATTTAAAGAAAGATTTAGAAA 480				
Qy 2000	ttcaaggtttaccagaaggttattcttactctgtccaagaacagatctcgaagactata 2059				
Db 481	TTTCAAACTTTTCAAGAAAAAACTTGATGAAGAAAGAGTTGAAMAAAGATCTCAGAGAAA 540				
Qy 2060	aggttaaggttaatagcccaagaagtgcgaatgctacagtttcaaaaacaggaataacaa 2119				
Db 541	AGAGAGAAAGAAATTTGAAAGAGAAATGAAGCACTTCMAAAGAAAAAAGAAATGTCGAAA 600				
Qy 2120	gtgattgagacactgtccttttgaataataaagagcgtgttctccacagagtgtac 2179				
Db 601	GTATGAAGATTTCTACGAGAAAGAGATTAAGACAACTTGACATTAATTTGAAAGACTACTC 660				
Qy 2180	aaaagatcaatcgtcatctagct 2202				
Db 661	AAAAGAAACTTGATGATATGACT 683				
RESULT 5					
CNS020K7	1092 bp DNA GSS 12-MAY-2000				
LOCUS	Tetradon nigriviridis genome survey sequence T7 end of clone				
DEFINITION	22211 of library G from Tetradon nigriviridis, genomic survey				
ACCESSION	sequence.				
VERSION	AL175696.1 GI:7813753				
KEYWORDS	GSS: genome survey sequence.				
SOURCE	Tetradon nigriviridis.				
ORGANISM	Tetradon nigriviridis.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Eurypterygii; Ctenosquamata; Acanthomorpha; Euacanthomorpha; Holacanthopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.				
AUTHORS	Roest-Crollius,H., Jallion,O., Dasilva,C., Filames,C., Fisher,C.,				

Db	513	ATTAAATAAATTTTAATATACAGTMAAAAAACCCTAATTTATATAAAAA	565
RESULT	7	CNS0DEPO	1101 bp DNA GSS 04-JUN-1999
CNS0DEPO/c		Drosophila melanogaster genome survey sequence T7 end of BAC:	
LOCUS		BACR23P01 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
DEFINITION		AL069493	
ACCESSION		AL069493.1 GI:4949636	
VERSION		GSS.	
KEYWORDS			
SOURCE		fruit fly.	
ORGANISM		Drosophila melanogaster	
REFERENCE		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS		1 (bases 1 to 1101)	
TITLE		Genoscope.	
JOURNAL		Direct Submissions Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.	
FEATURES		location/Qualifiers	
source		1..1101	
		/organism="Drosophila melanogaster"	
		/db_xref="taxon:7227"	
		/clone_1id="RPCI-98"	
		/clone="BACR29P01"	
		/note="end : 77"	
BASE COUNT		289 a 155 c 118 g 307 t 232 others	
ORIGIN			
Query Match		2.6% Score 60; DB 229; Length 1101;	
Best Local Similarity		33.0%; Pred. No. 0.0034;	
Matches 190; Conservative 94; Mismatches 291; Indels 1; Gaps 1;			
QY	1581	taaacataaaagactcattggttttggagaatgaatgatagtaacttagtcagtctaa	1640
DB	1030	DMRTGTTATRRRTGGTTAKTKWATRARMWTADGATGTAGTIDGTTTTTRTTRKDKAMA	971
QY	1641	aatccctgttagaataacgcctcaagaatagaaaccctccaagctaactgacctgattctt	1700
DB	970	WWMTTITTKTWGTTWGWTTWTTAKDTRTTDDWTTTADKTAAATGAAGTGAAMAAATTTMAMTW	911
QY	1701	tattccgatatcacataataatcatctcttatcgtaactcaagtgcataccagaagatt	1760
DB	910	MATAAKATTTFRAGAARTRTKTGTRRTTRDTTDRKAATTTTGATRGAGAKAGATTTMTA	851
QY	1761	agtgatatacttcgtatgtagaataaaaagaagaagtataacctgtaactcataattaac	1820
DB	850	TGTAATATATDTAACMMATAAATAATTAAMAAARAAMWTTTMTATATAMAAAATTTRTTAAMAA	791
QY	1821	attgagaanaacggtgactgctgcttacgcggygaacgaactaaagttccacttggaaat	1880
DB	790	ARKWRITAAAAAATATATATTTTAAAAAAMWATTTTTTTATATAATATTTMAMTWTTMTT	731

Oy	1881	tgaattaaaataatragagaagatctgttcctaactgfttaaaacgagataaacaa	1940
Dd	730	TTTGATTAAMWMAAAAAAAM-TAKTRARATPAATTATWATWATATATATWATTTTTT	672
Oy	1941	cctcgaaattaagaatgyltaaagcaaccatlaattttaaacaibyggaagtllacact	2000
Dd	671	TTTTTATWMTAAADWMTTPTATAATATATATATATATATATATATATATATATAT	612
Oy	2001	tcaaagflltacccaagtgattctcaccctgcgaagaacagatllcgaagctataa	2060
Dd	611	TTAARDATTTAAATTWMAAMWTAMAATAAPAAAATTTWAGGAAATTTTWMTTATATATW	552
Oy	2061	ggttaaatgataatagccaaagtagcaaatgctacagttlcaaacaacagagataacaag	2120
Dd	551	TTKTAATTAATAATTTWMTATTAATTAATWATWATTTTTTTTTTTTTTTTATTAATAATN	492
Oy	2121	tgatgagacactggctlltgaataataataaagacc	2156
Dd	491	AACCAANNAANNNGCTCCGCCCNCTTTTAANCCCC	456

RESULT	8
LOCUS	CNS0201P
DEFINITION	CNS0201P 1032 bp DNA GSS 14-MAY-2000 Tetradon nigroviridis genome survey sequence PUC-Orig end of clone 15N05 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION	AL206746.1 GI:7865565
KEYWORDS	GSS; genome survey sequence. Tetradon nigroviridis.
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Holoacanthopterygii; Acanthopterygii; Percomorpha; Holacanthopterygii; Tetraodontiformes; Tetraodontidae; Tetradon.
REFERENCE	1 (bases 1 to 1032) Roest-Crollius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C., Bonneau,L., Billault,A., Quetler,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
TITLE	Unpublished
JOURNAL	2 (bases 1 to 1032)
REFERENCE	RoeSt-CrOllIus,H., Jalllon,O., DasIlva,C., Bouneau,L., Fisher,C., BernOt,A., FItzameS,C., WInckler,P., Bröttler,P., Quetler,F., SaurIn,W. and WeIssenbach,J. Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
AUTHORS	Unpublished
TITLE	3 (bases 1 to 1032)
JOURNAL	Genoscope.
AUTHORS	Direct Submission
COMMENT	Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetradon .
FEATURES	Location/Qualifiers
SOURCE	1..1032 /organism="Tetradon nigroviridis" /db_xref="taxon:99883" /clone="15N05" /clone_1bp="G" /note="Genoscope sequence ID : C0XG15CG03SP1-end ; PUC-Orig"
BASE COUNT	489 a 103 c 100 g 284 t 56 others
ORIGIN	
Query Match	2.6% Score 59.2; DB 230; Length 1032;
Best Local Similarity	41.5%; Pred. No. 0.0051;

Best Local Similarity 36.4%; Pred. No. 0.0089;
Matches 205; Conservative 67; Mismatches 287; Indels 4; Gaps 2;

```

Qy 1569 attagaataagataaactaaagactatcagtgcttgagagacatga-atgatagtaact 1627
    || || || || || || || || || || || || || || || || || || || || || ||
Db 616 ATGAATAAAGAAAAAATGACAGCTGAATTTGAACTGGAAGCAGCATCMTTAAAGTTT 557
Qy 1628 tagaagtgctaaatccctgtagaatacgcacagaatagtaatccctccacagtaactg 1687
    || || || || || || || || || || || || || || || || || || || || || ||
Db 556 TATATAAGGGGCTGTGTGCTGAAAGTAAAMAAATGTTTATTAAMAAATCTGTGGG 497
Qy 1688 acctgattcttattatccgaataaataatcaatccctattggaactcagtgagc 1747
    || || || || || || || || || || || || || || || || || || || || || ||
Db 496 RTTWTGTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 437
Qy 1748 atccagaagattagtgatattatcgtatgagagataaaagaagtataactgttaa 1807
    || || || || || || || || || || || || || || || || || || || || || ||
Db 436 NNNAATTTNNNNNNNNNAAGCTTWARRRMAAATAAAMAAARRAAMAAATAAAMAA 377
Qy 1808 ctcaatacttaacatcgagaaacggtgactggttagctgtgtgacagaactaagatt 1867
    || || || || || || || || || || || || || || || || || || || || || ||
Db 376 WAAAAAATAAAGRCWAAATAATARMMMTAAGGACCTTTAAMAAAAAATAAAMAA 317
Qy 1868 tccatttgaattgaattaaataataagaagaattgcttctccaaactgttaaa 1927
    || || || || || || || || || || || || || || || || || || || || || ||
Db 316 AATCWRRTATTAAMMTTWTMAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 257
Qy 1928 cagataaacaacactcgaatttaagaatgtaagaacacatatttaaaacatgag 1987
    || || || || || || || || || || || || || || || || || || || || || ||
Db 256 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 197
Qy 1988 aaagtttaacacatcgaggtgtaccagaaggttaacttactgttcaagaacagact 2047
    || || || || || || || || || || || || || || || || || || || || || ||
Db 196 WAWMTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 137
Qy 2048 ctgagaagctaaagtttaagtaagccagaagatgacaaatgcacagttccaaaaa 2107
    || || || || || || || || || || || || || || || || || || || || || ||
Db 136 TTCCAAATTTTCA---WAAATTTATATRAAAAAAATAAATAAATAAATAAATAAATAA 80
Qy 2108 caggaataacaagtgatcgagaca 2130
    || || || || || || || || || || || || || || || || || || || || || ||
Db 79 AAAAAAATACMTWTWAMCAAMWA 57

RESULT 11
AZ683582 891 bp DNA GSS 14-DEC-2000
LOCUS ENTXK477r Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic DNA sequence.
ACCESSION AF683582
VERSION AF683582.1 GI:11820728
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 891)
AUTHORS Loftus,B., Van Aken,S. and Fraser,C.
TITLE Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9112 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 0208
Email: b1loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: Shotgun
High quality sequence start: 16
High quality sequence stop: 694.

```

FEATURES
source

Location/Qualifiers
1..891
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica sheared DNA"
/note="Vector: PHOS1, Site-1. Bst I. Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + 1 method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaundin and B.
Barell, Oxford University Press, 1999)."

BASE COUNT 400 a 55 c 291 g 145 t
ORIGIN

Query Match 2.5%; Score 57.8; DB 225; Length 891;
Best Local Similarity 43.1%; Pred. No. 0.01;
Matches 278; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

```

Qy 1547 attattcactgtagtgcgtgaattagaataagaactaaagactatcagtttg 1606
    || || || || || || || || || || || || || || || || || || || || || ||
Db 39 ATGATGAAGATGATTAATGATCAACATGATGATTAATTTGAATGAAGACGATGATGAAGATG 98
Qy 1607 gagacatgaatgataagtaacttagcaggtgcataaatccctgtgaatacgcctaagata 1666
    || || || || || || || || || || || || || || || || || || || || || ||
Db 99 ATGATGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 158
Qy 1667 gtaatccccaagctaacctgacttgattcttattccgaataaataataataat 1726
    || || || || || || || || || || || || || || || || || || || || || ||
Db 159 ATGATGATGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 218
Qy 1727 ctctattggaactcagtgagcatccagaagaattagttgatatattcgtatgagaata 1786
    || || || || || || || || || || || || || || || || || || || || || ||
Db 219 ATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 278
Qy 1787 aaaaagaagtataacctgttaactataatlaacatgagaataaagcggtgactgttag 1846
    || || || || || || || || || || || || || || || || || || || || || ||
Db 279 AAGAAAGACGATGAAGAAGACGATGATGATGATGATGATGATGATGATGATGATGATG 338
Qy 1847 ctgttgacagaactaaagattccatttgaattgaataaataataataaagaagaat 1906
    || || || || || || || || || || || || || || || || || || || || || ||
Db 339 AAGAAAGACGATGAAGAAGACGATGATGATGATGATGATGATGATGATGATGATGATG 398
Qy 1907 tgccttcacaactgtttaaacaagataaacaaccccgaaatttaagaatgtaagaacaa 1966
    || || || || || || || || || || || || || || || || || || || || || ||
Db 399 ACGATGATGAAGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 458
Qy 1967 ccattaatlaaacaatgagggaagtttaaacacttcaaggtttaaccagaaggttaacct 2026
    || || || || || || || || || || || || || || || || || || || || || ||
Db 459 ACGATGATGAAGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 518
Qy 2027 acctgtcagaagaacagattcgaagctataaggttaagtttaagccaagaagtaag 2086
    || || || || || || || || || || || || || || || || || || || || || ||
Db 519 ACGATGATGAAGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 578
Qy 2087 caaatgttaagtttcaaaaacaggaataaacaagtatgagacactgtcttttgaataata 2146
    || || || || || || || || || || || || || || || || || || || || || ||
Db 579 ACGATGATGAAGAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 638
Qy 2147 ataagaagcctgtgttctccacagagagttgatcaaaagaatcaatg 2191
    || || || || || || || || || || || || || || || || || || || || || ||
Db 639 AAGACGATGAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 683

```

RESULT 12

LOCUS	CNS03HOU	652 bp	DNA	GSS	17-MAY-2000
DEFINITION	Tetraodon nigroviridis genome survey sequence T7 end of clone 02K10 of library G from Tetraodon nigroviridis, genomic survey sequence.				
ACCESSION	AL244551.1 GI:7965563				
VERSION	AL244551				
KEYWORDS	GSS; genome survey sequence.				
SOURCE	Tetraodon nigroviridis.				
ORGANISM	Tetraodon nigroviridis				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Eurypterygii; Ctenosquamata; Acanthomorpha; Eucanthomorpha; Holacanthopterygii; Acanthopterygii; Petcomorpha; Tetraodontiformes; Tetraodontoidae; Tetraodontidae; Tetraodon.				
AUTHORS	1 (bases 1 to 652) Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.				
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 652) Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Queller,F., Saurin,W. and Weissenbach,J.				
AUTHORS	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence				
TITLE	Unpublished				
JOURNAL	3 (bases 1 to 652) Genoscope.				
REFERENCE	Direct Submission				
AUTHORS	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases				
TITLE	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .				
COMMENT	Location/Qualifiers				
FEATURES	1..652				
SOURCE	/organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="02K10" /clone.lib="G" /note="Genoscope sequence ID : C08G027BF05LP1-end : 'T7'"				
BASE COUNT	385 a 51 c 41 g 73 t 102 others				
ORIGIN					
Query Match	2.5%; Score 57.4; DB 231; Length 652;				
Best Local Similarity	38.4%; Pred. No. 0.012;				
Matches 178; Conservative	4; Mismatches 282; Indels 0; Gaps 0;				
OY	1535 agttgacatlatatttcctgtagctgtagctggaattgataaggaataactaaagact 1594				
Db	173 AANTNNAAANWGTAAATTATCATANAAGCAANNAAAAAANAAAAAANAAAAANA 232				
OY	1595 atcagtgtttgagacatgatagtacttagctgacgtgctaaatccctgtgagat 1654				
Db	233 AAMNMAAAAAAANCAAAAAAANAAATANNNNANNNNNNGNAAANNAGTGAAGAAG 292				
OY	1655 acgctcaagatagtaatccctccacagcgaactgactgattcttcttccgaaataca 1714				
Db	293 NNGAAMAAATTAANNNNNNNNAAAAANNTTTTAAANTNNNATATNATRCRAAATAA 352				
OY	1715 ataataatcatctcttttggactcgtgcatccgagaatttgatgatattatc 1774				
Db	353 AAAAAAAATATTTAANTANNAANNNNNNNNNAAAAATTAATNAANTTAATTAATAA 412				
OY	1775 gtatcgagatcaaaaaaagatctatcctgtaactcacaattaatcattgagaacagc 1834				
Db	413 ANMAAAAAAGAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAANA 472				
OY	1835 tgactggtttagctggtgacgaactaaagattccatttgaattgaatgaataaata 1894				

[illegible]

Oy	1874	tctgaatgcatcctaataatlaagcaagaattgccttccaaactgttaaacacagta	1933
Db	899	WAAAAATGAAAATAATTAAWTMTTAAATTTTAATTTTTTTTAATTAACAAAAAATTAACAAA	840
Oy	1934	aacacaacctcgcaatttaagaagtgaagcacaccatlaattaaaacatggggaagtt	1993
Db	839	AAATTAATTAACAAAATAAAAMMAAAWMATMMAAAVNTAAMAATAATMAAAAAAHATAATMAAAAT	780
Oy	1994	taacactcaagggttaccagaaggttccttcttacctgtcaaagaacagattctgaag	2053
Db	779	ACMAAAATYMATTTTAAAAATTAAMWAAMWTMTWAAMWMMNATATMAWMAAAAAATTMAATT	720
Oy	2054	gctataaggttaaagtgatatagccagaagtagcacaatgctacagttccaacacaggaa	2113
Db	719	HCMMAATTAACAAACMAAAHAAHAAAMCMAAAAAAATTTTCCCAAWATCACAIYAACCNA	660
Oy	2114	taacaagtatgagacactgtccttttgaanaataaagaagcctgttgcctacaggag	2173
Db	659	CACAATAATCAAAACHAACACAAIYAAYAACMAAAVCACAATATBCAATATCAIYAAGRSDCY	600
Oy	2174	tgtatcaaaagaatcaat	2190
Db	599	YAAATYAAATKSYHHAW	583
RESULT	14		
CNS00KAE			
LOCUS	CNS00KAE	1101 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR39P05 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	AL077628	GI:4956903
ACCESSION	AL077628		
VERSION	1		
KEYWORDS	GSS.		
SOURCE	Fruit fly.		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
AUTHORS	JOURNAL		
TITLE	Genoscope.		
COMMENT	Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm. Location/Qualifiers 1..1101 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone_lib="RPCI-98" /clone="BACR39P05" /note="end : TET3"		
FEATURES	source		
BASE COUNT	374 a	59 c	46 g
ORIGIN		399 t	223 others
Query Match	2.5%	Score 56.8;	DB 229; Length 1101;

QY	1425	tgcaccttccttaaacacatccaaagaagctatgagaagggttaacagggaagaagaca	1484
Db	140	TNANANAMNNNTTATTAATATATATNANTATANTNATNANANANANATATATANANNTNT	199
QY	1485	agccattgagctatgctgtcttaacgcgaacacaaatctgcggtactactgctgaacat	1544
Db	200	ANTNTATGMCANMWNNTATATANATATNTNTNNTNMTNATATATATATAMNNAN	259
QY	1545	atatattccactgcagtagctgcgaattgataaagata--actaaagactcatgct	1602
Db	260	ANATATANANNTNAAATANANANATATGCMGATATATATATATAAAAAATATATATA	319
QY	1603	tttggagacatgaatgatacttactttagcagcttgcataaactctgtagaatgcctca	1662
Db	320	ATATATATATATATATATATATATATATATATNTWTWAAAMAAAAATWTATTTATATNA	379
QY	1663	gatgatactccaccagcagctaacgaccttgattcttcttactccgaataacataaat	1722
Db	380	AAAAATTTTATATATTTAAATATATATWAAATAAAAATTAATATATATATATATATAA	439
QY	1723	caactcttatctggagaccagctgcagccacagaagaattagctgatactatccgatgaa	1782
Db	440	TWAAATATATATTTAAATATATATATATATTTTAMAMAAITTTAATATATATWAAATAAA	499
QY	1783	gataaaaaagaagtataccctgtaactcataatattaacatctgagaagaacgctgctgt	1842
Db	500	WATATAMAMAMAAATANAAATATATATTTTAAATTAAMAAATTAATAMAMAAATATAN	554
QY	1843	tttagctgtgcagacgaactaaagattccatttgcgaattgaaatgaaataataagcaa	1902
Db	555	--AAWTTTATTAAMWMAAAWMAAMAAATATMAAAAAATTAAMAAAAATATATATA	612
QY	1903	gaatgcttctcaaacgcttaaacagataaacaacacccgaattttaaagatgtaaa	1962
Db	613	AAAMAAATTAATATATTTAAATTAAMAAAAATATAAAAAATATATATATAAAAAATATA	672
QY	1963	gcaacattatcttaaaacat	1983
Db	673	AAAAGSTTTTACTTCGRMCTT	693
RESULT	15		
BF274633			
LOCUS			
DEFINITION	GA_EB0021J19f Gossypium arboreum 7-10 dpa fiber library Gossypium	EST	17-NOV-2000
ACCESSION	GA_EB0021J19f Gossypium arboreum 7-10 dpa fiber library Gossypium		
VERSION	BF274633		
KEYWORDS	BF274633.1 GI:11205703		
SOURCE	EST.		
ORGANISM	Gossypium arboreum.		
REFERENCE	Gossypium arboreum.		
AUTHORS	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvaceae; Gossypium.		
TITLE	1 (bases 1 to 928)		
JOURNAL	Wang R.A., Flischi D., Yu Y., Main D., Rambo T., Simmons J., Henry		
COMMENT	'D., Wood, T.C., Leslie, A. and Wilkins, T.A. An integrated analysis of the genetics, development, and evolution of the cotton fiber unpublished (2000) Contact: Ming RA Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 7288 Fax: 864 656 4293 Email: rtwing@clemson.edu Seq primer: TAATACGACTCATATAGG High quality sequence start: 105 High quality sequence stop: 96.		

FEATURES

Source

Location/Qualifiers
1 928

1. .928

```
/organism="Gossypium arboreum"
```

```
/scli all= ana
/cu]tivar="8400"
```

```

/cutcutcut - 0400
/dh yref="taron:

```

```
/usr/src/.../clone="GA Eb0021J19f
```

```
clone_lib="Gossypium a
```

/tissue_type="Fibers isolated from bolls harvested 7-10

дпа"

```
/lab_host="E. coli"
```

```
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
```

Category	Count
a	85
c	80
g	248
t	15
others	

ORIGIN

Query Match

Best Local Similarity 47.1%; Pred. NO. 0.021;

matches	1/1;	conservative	0;	mismatches	1
---------	------	--------------	----	------------	---

QY 1770 tattcgtatggaagataaaaaagaagtataccctgtaactcataatlttaacattgagaaa 1825

Db 406 TATAAGATTCAAATACAAAAAATTAGGCCAGAAAAATAAAGATAAAAAATGAAAA 465

1830 aacggtgactggtttagctggtgacagaactaaagatttccatttgaattgaattaa 1885

Db 466 ACACCTAGATTACCCAATTAAATCAATATATAAAAAAATCACCCTATACAAAAATTAANAAA 525

1890 aaataagaagaattgtcttctcaactgttaaacagataaacacaactcgaatt 1945

Db 526 AAAAAAAAAACACAAAAATACTAAACAAAAATAAATTAACATAAAAAA 585

OY 1950 taaagatggttaaagcaaccattaatttaaaacatggggaagtttaacacttcaagttt 2005

[illegible]

ov. 2010 accaagaaatattcttacctgtcaagaacacaaatctgaagctataaacttaaat 2065

[illegible]

0v 3070 taaatacccccaaatatccaaaatatcactttcaaaaaacgaaataaaccaattgatgaan 3136

[illegible]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

QY	2130	act	2132

Search completed: June 6, 2001, 20:03:01
Job time: 4842 sec